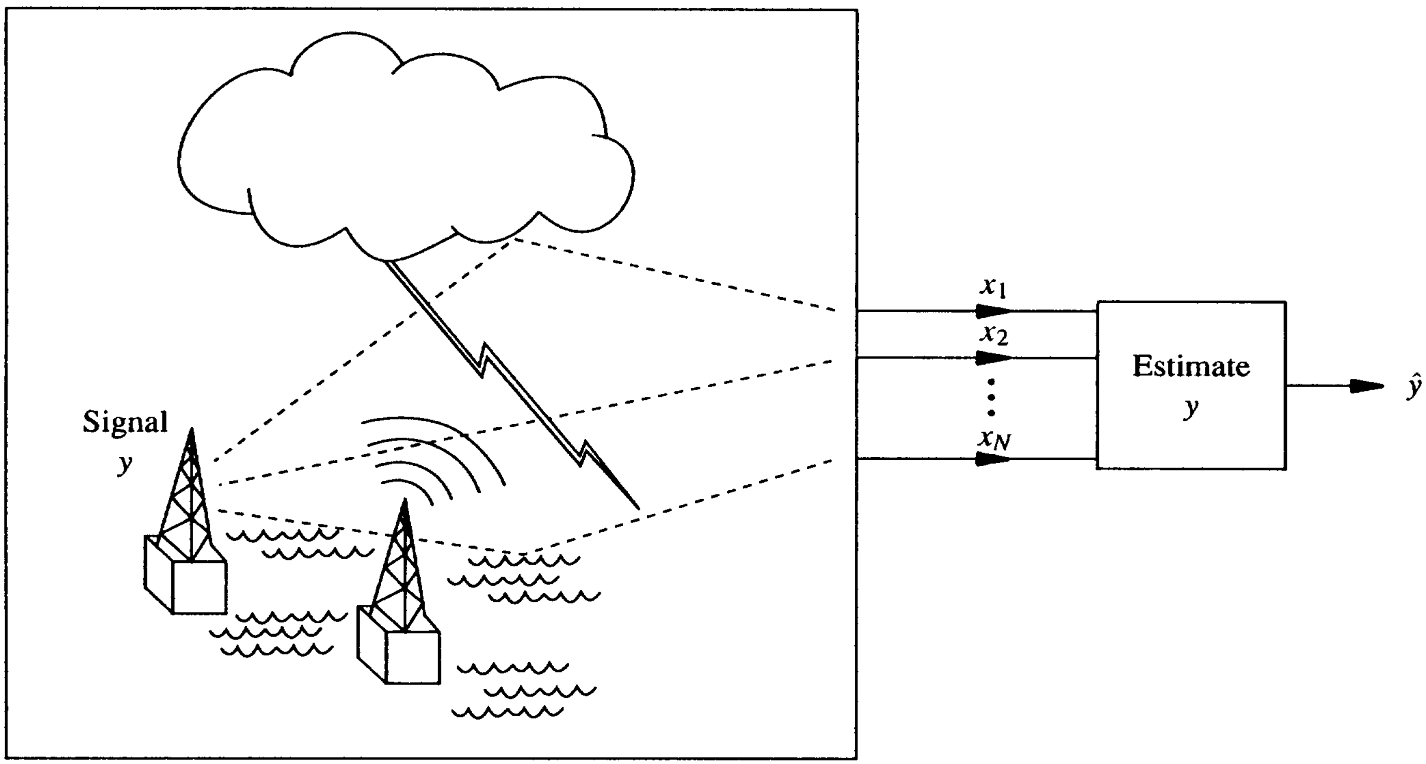
ESTIMATION PROBLEM



y: a random variable
$x_1, \cdots, x_N$: a set of related measurements.

LMS ESTIMATION REVISITED

FORM OF ESTIMATE

$$\widehat{y} = \mathbf{a}^{*T}\boldsymbol{x} = a_1^* x_1 + a_2^* x_2 + \ldots + a_N^* x_N$$

ORTHOGONALITY PRINCIPLE

Theorem: Let $\varepsilon = y - \widehat{y}$ be the error in estimation. Then $\mathbf{a}$ minimizes the mean-square error $\sigma_\varepsilon^2 = \mathcal{E}\left\{|y - \widehat{y}|^2\right\}$ if $\mathbf{a}$ is chosen such that $\mathcal{E}\{x_i \varepsilon^*\} = \mathcal{E}\left\{\varepsilon x_i^*\right\} = 0 \quad i = 1, 2, \ldots, N$, that is, if the error is orthogonal to the observations. Further, the minimum mean-square error is given by $\sigma_\varepsilon^2 = \mathcal{E}\{y\varepsilon^*\} = \mathcal{E}\{\varepsilon y^*\}$.

PROOF OF ORTHOGONALITY

Let $\mathbf{a}$ be any weighting vector and $\mathbf{a}^{\perp}$ be the weighting vector that results in orthogonality. Then

$$\varepsilon = y - \mathbf{a}^{*T}\boldsymbol{x} = y - \left(\mathbf{a}^{\perp}\right)^{*T}\boldsymbol{x} + \left(\mathbf{a}^{\perp} - \mathbf{a}\right)^{*T}\boldsymbol{x} = \varepsilon^{\perp} + \left(\mathbf{a}^{\perp} - \mathbf{a}\right)^{*T}\boldsymbol{x}$$

where $\varepsilon^{\perp}$ is the error that results from using $\mathbf{a}^{\perp}$. Now observe

$$\begin{aligned}
\sigma_{\varepsilon}^{2} = \mathcal{E}\left\{|\varepsilon|^{2}\right\} &= \mathcal{E}\left\{(\varepsilon^{\perp} + \left(\mathbf{a}^{\perp} - \mathbf{a}\right)^{*T}\boldsymbol{x})(\varepsilon^{\perp} + \left(\mathbf{a}^{\perp} - \mathbf{a}\right)^{*T}\boldsymbol{x})^{*}\right\} \\
&= \mathcal{E}\left\{|\varepsilon^{\perp}|^{2}\right\} + \left(\mathbf{a}^{\perp} - \mathbf{a}\right)^{*T}\mathcal{E}\left\{\boldsymbol{x}(\varepsilon^{\perp})^{*}\right\} \\
&\qquad + \mathcal{E}\left\{\varepsilon^{\perp}\boldsymbol{x}^{*T}\right\}\left(\mathbf{a}^{\perp} - \mathbf{a}\right) + \mathcal{E}\left\{(|\mathbf{a}^{\perp} - \mathbf{a})^{*T}\boldsymbol{x}|^{2}\right\} \\
&= \mathcal{E}\left\{|\varepsilon^{\perp}|^{2}\right\} + \mathcal{E}\left\{|(\mathbf{a}^{\perp} - \mathbf{a})^{*T}\boldsymbol{x}|^{2}\right\}
\end{aligned}$$

This is minimized when $\mathbf{a} = \mathbf{a}^{\perp}$.

PROOF OF ORTHOGONALITY (cont'd.)

To find the minimum mean-square error:

$$
\begin{aligned}
(\sigma_\varepsilon^2)_{MIN} = \mathcal{E}\left\{\varepsilon^\perp(\varepsilon^\perp)^*\right\} &= \mathcal{E}\left\{(y - (\mathbf{a}^\perp)^{*T}\boldsymbol{x})(\varepsilon^\perp)^*\right\} \\
&= \mathcal{E}\left\{y(\varepsilon^\perp)^*\right\}
\end{aligned}
$$

LMS ESTIMATION PROBLEM

- Random variable y, vector of observations $\boldsymbol{x}$, both have *zero mean* and may be complex.

- Estimate of the form: $\widehat{y} = \mathbf{a}^{*T}\boldsymbol{x}$ where

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

- Choose $\mathbf{a}$ to minimize $\varepsilon_{lms} = \mathcal{E}\left\{|y - \widehat{y}|^2\right\}$

SOLUTION TO LMS ESTIMATION

Apply orthogonality principle:

$$\mathcal{E}\{\boldsymbol{x}\varepsilon^*\} = \mathcal{E}\left\{\boldsymbol{x}(y - \mathbf{a}^{*T}\boldsymbol{x})^*\right\} = \mathcal{E}\left\{\boldsymbol{x}(y^* - \boldsymbol{x}^{*T}\mathbf{a})\right\} = \mathbf{0}$$

Thus

$$\mathbf{r}_{\boldsymbol{x}y} - \mathbf{R}_{\boldsymbol{x}}\mathbf{a} = 0 \qquad \text{or...} \qquad \boxed{\mathbf{R}_{\boldsymbol{x}}\mathbf{a} = \mathbf{r}_{\boldsymbol{x}y}}$$

The mean-square error is:

$$\sigma_\varepsilon^2 = \mathcal{E}\{y\varepsilon^*\} = \mathcal{E}\left\{y(y^* - \boldsymbol{x}^{*T}\mathbf{a})\right\}$$

$$\text{or ...} \qquad \boxed{\sigma_\varepsilon^2 = \sigma_y^2 - \mathbf{r}_{\boldsymbol{x}y}^{*T}\mathbf{a}}$$

POSTULATES FOR A VECTOR SPACE

1. A *vector space* $\mathcal{V}$ is a set of elements $u, v, \cdots$ such that if $u \in \mathcal{V}$ and $v \in \sqsubseteq$ then there is a unique element

$$u + v \in \mathcal{V}$$

called the *sum*. Further if c is from an associated field, such as the field of complex numbers, then the *scalar product* is an element

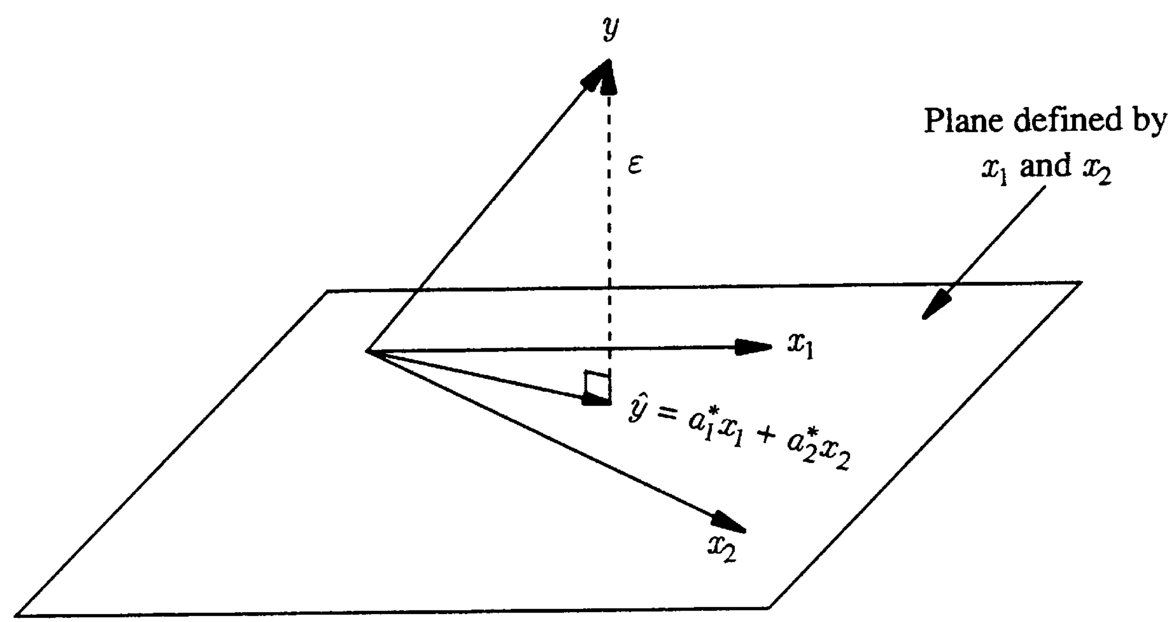
$$c \cdot u \in \mathcal{V}$$

with certain associative and distributive properties.

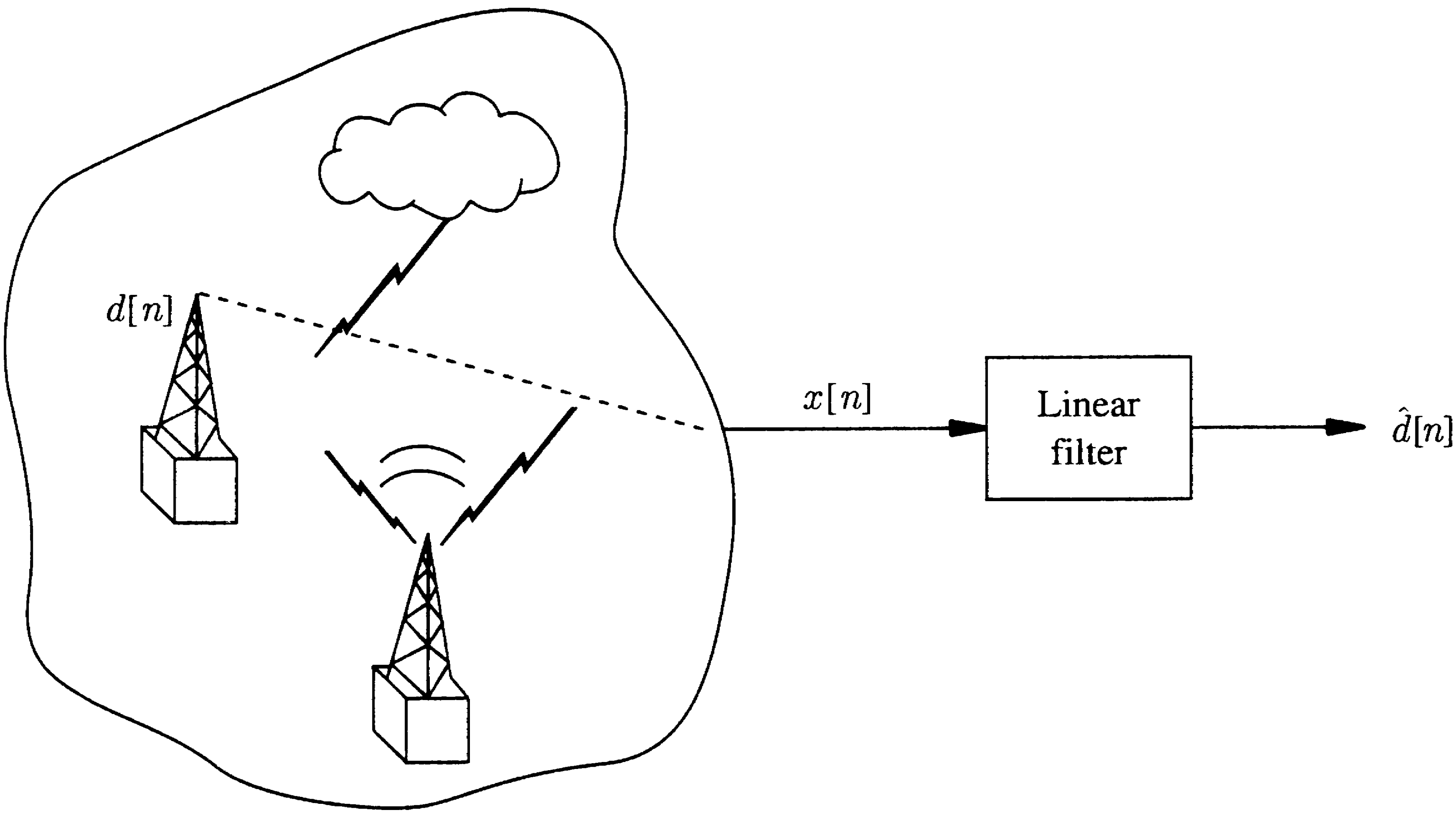
2. A vector space is an inner product space or a *Hilbert space* if an inner product (u, v) between elements is defined.

VECTOR SPACE INTERPRETATION OF THE ORTHOGONALITY PRINCIPLE

- Elements of $\mathcal{V}$ are random variables.
- Inner product is expectation; e.g., $(y, \varepsilon) \stackrel{\text{def}}{=} \mathcal{E}\{y\varepsilon^*\}$.



OPTIMAL FILTERING PROBLEM



TYPICAL WIENER FILTERING PROBLEMS

Problem	Form of Observations	Desired Sequence
Filtering: signal in noise	$x[n] = s[n] + \eta[n]$	$d[n] = s[n]$
Prediction: signal in noise	$x[n] = s[n] + \eta[n]$	$d[n] = s[n+p]$
Smoothing: signal in noise	$x[n] = s[n] + \eta[n]$	$d[n] = s[n-q]$
Linear prediction	$x[n] = s[n-1]$	$d[n] = s[n]$
General nonlinear problem	$x[n] = G\left(s[n], \eta[n]\right)$	$d[n] = s[n]$

CONCEPTS FOR OPTIMAL FILTERING

FORM OF ESTIMATE AND ERROR

$$\hat{d}[n] = (\tilde{\boldsymbol{x}}[n])^T \mathbf{h}\,; \qquad \varepsilon[n] = d[n] - \hat{d}[n]$$

where $\quad \tilde{\boldsymbol{x}}[n] = \begin{bmatrix} x[n] \\ x[n-1] \\ \vdots \\ x[n-P+1] \end{bmatrix} \qquad \mathbf{h} = \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[P-1] \end{bmatrix}$

ORTHOGONALITY PRINCIPLE

$$\mathcal{E}\{\tilde{\boldsymbol{x}}[n]\varepsilon^*[n]\} = \mathbf{0}\,; \qquad \sigma_\varepsilon^2 = \mathcal{E}\{d[n]\varepsilon^*[n]\}$$

DERIVATION OF EQUATIONS

The orthogonality principle states:

$$\mathcal{E}\{\tilde{\boldsymbol{x}}[n]\varepsilon^*[n]\} = \mathcal{E}\{\tilde{\boldsymbol{x}}[n](d^*[n] - (\tilde{\boldsymbol{x}}[n])^{*T}\mathbf{h}^*)\} = \mathbf{0}$$

or $$\boxed{\mathbf{R}_{\boldsymbol{x}}\,\mathbf{h} = \tilde{\mathbf{r}}_{d\boldsymbol{x}}}$$ (Wiener-Hopf equation)

The minimum mean-square error is given by:

$$\sigma_\varepsilon^2 = \mathcal{E}\{d[n]\varepsilon^*[n]\} = \mathcal{E}\{d[n](d^*[n] - (\tilde{\boldsymbol{x}}[n])^{*T}\mathbf{h}^*)\}$$

or $$\boxed{\sigma_\varepsilon^2 = R_d[0] - \mathbf{h}^{*T}\tilde{\mathbf{r}}_{d\boldsymbol{x}}}$$

where $\mathbf{R}_{\boldsymbol{x}} = \mathcal{E}\left\{\boldsymbol{x}[n]\boldsymbol{x}^{*T}[n]\right\}$; $\mathbf{r}_{d\boldsymbol{x}} = \mathcal{E}\{d[n](\boldsymbol{x}[n])^*\}$

EQUATIONS FOR OPTIMAL FILTERING

WIENER-HOPF EQUATION

$$\begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-P+1] \\ R_x[1] & R_x[0] & \cdots & R_x[-P+2] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P-1] & R_x[P-2] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[P-1] \end{bmatrix} = \begin{bmatrix} R_{dx}[0] \\ R_{dx}[1] \\ \vdots \\ R_{dx}[P-1] \end{bmatrix}$$

MEAN-SQUARE ERROR

$$\sigma_\varepsilon^2 = R_d[0] - \sum_{l=0}^{P-1} h^*[l] R_{dx}[l]$$

Insert Example 7.2 here.

Insert Example 7.3 here.